|  |
| --- |
| **LEES EERST DIT** |
| Hierna vind je een synthese oefening. Dit is een ultieme rekentest, waarbij je uiterste concentratie wordt gevraagd.  **Het databestand van deze oefening is het format waarin de examenvragen zullen gepresenteerd worden.**  **Het is belangrijk dat je je hiermee vertrouwd maakt.**  De rijen in het databestand vertegenwoordigen statistische eenheden, de kolommen vertegenwoordigen de variabelen. Elke cel geeft voor één statistische eenheid de waarde weer op de betreffende variabele.  Op het examen worden meerdere vragen gesteld op basis van een analyse met gegevens afkomstig uit één databestand (in het format zoals hier in deze synthese oefening gepresenteerd). We testen je kennis over de univariate en bivariate statistiek dan op basis van één dataset, in plaats van op basis van verschillende datasets. Dat geeft minder rekenwerk. Let wel op dat fouten kunnen accumuleren, maar dat is altijd zo wanneer statistische parameters moeten worden berekend die samengesteld zijn uit verschillende tussenstappen.  We weten uit ervaring dat studenten moeite hebben met deze synthese oefening omdat alle leerstof samengebracht wordt. Studenten hebben de neiging om de leerstof in te studeren ‘in aparte hoofdstukken’. Wanneer alles gesynthetiseerd wordt in één grote oefening, dan loopt het vaak mis. Vandaar dat we ten zeerste aanraden om deze synthese oefening te maken als voorbereiding op het examen.  De oplossingen worden hier nog niet meegegeven en dit om didactische redenen, maar worden in de loop van de lessenreeks ter beschikking gesteld via Ufora. Meer informatie hieromtrent wordt later meegedeeld.  VEEL SUCCES!! |

**SYNTHESE-OEFENING**

**Sommige wetenschappers zijn van mening dat de kwaliteit van de relatie tussen adolescent en ouders een belangrijke factor in de relatie tot van huis weglopen. Op basis van een kleine toevalssteekproef van adolescenten werden de volgende data verzameld met betrekking tot sexe, het aantal uur dat respondenten wekelijks kwaliteitstijd doorbrengen met hun ouders (variabele X1), het aantal keren dat zij in de afgelopen zes maand overwogen om thuis weg te lopen (variabele X2), en studieresultaten (variabele Y).**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Adolescent** | **Sexe** | **X1**  **Uren kwaliteitstijd met de ouders** | **X2**  **Aantal keren overwogen om thuis weg te lopen** | **Y**  **Studieresultaten** |
| **1** | **Meisje** | **1** | **10** | **5** |
| **2** | **Meisje** | **10** | **2** | **13** |
| **3** | **Meisje** | **5** | **7** | **7** |
| **4** | **Meisje** | **7** | **1** | **16** |
| **5** | **Meisje** | **20** | **0** | **11** |
| **6** | **Meisje** | **12** | **2** | **17** |
| **7** | **Meisje** | **5** | **1** | **15** |
| **8** | **Meisje** | **19** | **0** | **9** |
| **9** | **Jongen** | **4** | **6** | **7** |
| **10** | **Jongen** | **10** | **1** | **16** |
| **11** | **Jongen** | **12** | **1** | **13** |
| **12** | **Jongen** | **4** | **0** | **5** |
| **13** | **Jongen** | **10** | **0** | **12** |
| **14** | **Jongen** | **6** | **1** | **9** |
| **15** | **Jongen** | **2** | **9** | **5** |
| **16** | **Jongen** | **9** | **1** | **11** |

* Bereken de gemiddelde waarde voor de variabelen X1, X2 en Y.
* Wat is het gemiddelde van de jongens op X1, X2 en Y?
* Wat is het gemiddelde van de meisjes op X1, X2 en Y?
* Verschillen jongens en meisjes statistisch significant op X1, X2 en Y?
* Wat is de variatie, variantie, standaardafwijking voor X1, X2 en Y?
* Bereken de variatie, variantie, standaardafwijking afzonderlijk voor de jongens en de meisjes
* Wat is de covariatie, covariantie en correlatie tussen X1 en Y, tussen X2 en Y, tussen X1 en X2?
* Zijn deze bivariate associatiematen gelijk voor de jongens als voor de meisjes?
* Bereken de parameters van de regressieanalyse voor de regressie van Y op X1, en voor de regressie van X1 op Y.
* Bereken de correlatie tussen de verwachte waarde voor Y op basis van X1 en de geobserveerde waarde voor Y. Is deze waarde gelijk aan de correlatie tussen X1 en Y?
* Hoe groot is de determinatiecoëfficiënt voor de regressie van Y op basis van X1?
* Hoe groot is de determinatiecoëfficiënt voor de regressie van Y op basis van X2?
* Deze dataset is een toevalsteekproef. Bereken de betrouwbaarheidsintervallen voor de gemiddelde scores van X1, X2 en Y.
* Bereken de partiële correlatiecoëfficiënt tussen X1 en Y onder controle van X2.

Wat is jouw besluit ?

* Hoeveel van de variabiliteit in ‘studieresultaten’(Y) kan verklaard worden op basis van ‘uren kwaliteitstijd met de ouders’ (X1) en ‘aantal keren overwogen om thuis weg te lopen’ (X2) ?
* Welke factor (X1 of X2) heeft relatief het sterkste effect op studieresultaten ?
* Wat is het verwachte studieresultaat voor een adolescent die wekelijks 15u kwaliteitstijd met de ouders doorbrengt en de afgelopen zes maanden geen enkele keer heeft overwogen om thuis weg te lopen ?